```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          March 14, 2002, 17:23:11 ; Search time 76.91 Seconds
                                            (without alignments)
                                            12.521 Million cell updates/sec

Title:           US-09-786-009-7
Perfect score:   69
Sequence:        1 MAMGGGRLEGSSC 13

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters:       522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_1101:*
                1:  /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                              Description
-----------------------------------------------------------------------------
   1       43    62.3    371  20  AAY06367                       Streptomyces sp. E
   2       43    62.3    371  21  AAB14889                       Streptomyces livid
   3       43    62.3    371  21  AAY84345                       Amino acid sequenc
   4       43    62.3    386  21  AAY67497                       Protien sequence o
   5       40    58.0     33  21  AAB29255                       Gene 17 human secr
   6       40    58.0     54  21  AAG08827                       Arabidopsis thalia
   7       40    58.0    432  20  AAW88236                       Mouse prothrombina
   8       39    56.5     25  20  AAY36511                       Fragment of human
   9       39    56.5    106  22  AAB95038                       Human protein sequ
  10       39    56.5    647  22  AAY97667                       Zea mays ZmEIN3 2
  11       38    55.1    137  21  AAG07260                       Arabidopsis thalia
  12       38    55.1    147  2[illegible]  AAY59[illegible]4                   Human normal [illegible]
  13       38    55.1    269  21  AAG44[illegible]8                   Arabidopsis thalia
  14       38    55.1    270  21  AAG52[illegible]                   Arabidopsis thalia
  15       38    55.1    318  22  AAM40912                       Human polypeptide
  16       38    55.1    334  21  AAG44[illegible]7                   Arabidopsis thalia
  17       38    55.1    335  21  AAG52[illegible]9                   Arabidopsis thalia
  18       38    55.1    344  21  AAG44[illegible]6                   Arabidopsis thalia
  19       38    55.1    345  21  AAG52[illegible]8                   Arabidopsis thalia
  20       38    55.1    [illegible]  2[illegible]  AAB[illegible]                   Phytyl/prenyltrans
  21       38    55.1    455  2[illegible]  AAY28[illegible]4                   Aspergillus [illegible]
  22       38    55.1    456  2[illegible]  AAY28[illegible]5                   Redesigned Asperg
  23       37    53.6    117  2[illegible]  AAG0[illegible]8                   Human secreted pro
  24       37    53.6    119  2[illegible]  AAB[illegible]                   Human [illegible]
  25       37    53.6    2[illegible]  2[illegible]  AAY[illegible]7                   Protein inv[illegible]
  26       37    53.6    3[illegible]  2[illegible]  AAG0[illegible]                   Arabidopsis thalia
  27       37    53.6    3[illegible]  2[illegible]  AAG1[illegible]                   Arabidopsis thalia
  28       37    53.6    3[illegible]  2[illegible]  AAG4[illegible]6                   Arabidopsis thalia
  29       37    53.6    309  2[illegible]  AAG0[illegible]8                   Arabidopsis thalia
  30       37    53.6    3[illegible]9  2[illegible]  AAG1[illegible]                   Arabidopsis thalia
  31       37    53.6    3[illegible]9  2[illegible]  AAG4[illegible]5                   Arabidopsis thalia
  32       37    53.6    3[illegible]  2[illegible]  AAY[illegible]                   [illegible]
  33       37    53.6    371  2[illegible]  AAY67[illegible]6                   Actinomycete [illegible]
  34       37    53.6    4[illegible]  1[illegible]  AAR3[illegible]5                   E. coli thy[illegible]
  35       37    53.6    44[illegible]  1[illegible]  AAR3[illegible]6                   E. coli thy[illegible]
  36       37    53.6    44[illegible]  2[illegible]  AAB6[illegible]3                   Amino acid sequenc
  37       37    53.6    672  2[illegible]  AAM3[illegible]7                   Human polypeptide
  38       37    53.6    7[illegible]9  2[illegible]  AAM3[illegible]6                   Human polypeptide
  39       37    53.6    746  2[illegible]  AAM4[illegible]2                   Human polypeptide
  40       37    53.6    746  2[illegible]  AAM4[illegible]3                   Human polypeptide
  41       36    52.2     28  2[illegible]  AAW9[illegible]43                   Human BS124 specif
  42       36    52.2    [illegible]  2[illegible]  AA[illegible]1                   Human secreted pro
  43       36    52.2    1[illegible]  2[illegible]  AAY31750                       Human CPRO[illegible] rela
  44       36    52.2    [illegible]  2[illegible]  [illegible]                   [illegible]
  45       36    52.2    1[illegible]  2[illegible]  AAW9[illegible]41                   Human BS124 specif


                              ALIGNMENTS


RESULT    1
AAY06367
ID   AAY06367 standard; Protein; 371 AA.
XX
AC   AAY06367;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Streptomyces sp. EGIII-like cellulase.
XX
KW   Cellulase; endoglucanase; EGIII; textile; feed additive; baking;
KW   food processing; grain wet milling; pulp; paper.
XX
OS   Streptomyces sp.
XX
PN   WO9931255-A2.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;    98WO-US26551.
XX
PR   16-DEC-1997;    97US-0991720.
XX
PA   (GEMV ) GENENCOR INT INC.
XX
PI   Bower BS,  Fowler T,  Phillips JH.
XX
DR   WPI; 1999-405187/34.
XX
PT   EGIII like cellulase
XX
PS   Example; Fig 6; 4pp; English.
XX
CC   The present polypeptide represents a [illegible] sequence of a
```

```
CC   novel EGIII like cellulase of Streptomyces sp. 11AG8.  It was
CC   deduced from a gene sequence isolated from genomic DNA using PCR
CC   primers (see AAX59180-91) based on conserved motifs (see AAY06325-29)
CC   of Trichoderma reesei EGIII cellulase and related enzymes.  PCR
CC   has been used to identify novel EGIII like enzymes, including the
CC   present protein, from bacterial and fungal sources (see AAY06331-70).
CC   The sequence shows homology to T. reesei EGIII (see AAY06330).  Also
CC   provided by the invention are vectors, host cells and methods
CC   for the recombinant production of such enzymes, which can be used
CC   in the treatment of cellulose containing textiles, as feed
CC   additives, in the treatment of wood pulp, in the reduction of
CC   biomass to glucose, in the stone washing of indigo dyed denim, or
CC   as laundry detergent components (all claimed).
XX
SQ   Sequence   371 AA;


  Query Match              62.3%;  Score 43;  DB 20;  Length 371;
  Best Local Similarity   66.7%;  Pred. No. 30;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

          2 AMGGGRLEGSSC 13
            |  | || |:|
Db      357 apaggrlnqtsc 368


RESULT   2
AAB14880
ID   AAB14880 standard; Protein; 371 AA.
XX
AC   AAB14880;
XX
DT   21 NOV 2000  (first entry)
XX
DE   Streptomyces lividans CelB EGIII like cellulase.
XX
KW   Streptomyces lividans; Trichoderma reesei; endoglucanase III; EGIII;
KW   cellulase; mutant; enzyme stability; textile treatment;
KW   wood pulp treatment; feed additive; detergent.
XX
OS   Streptomyces lividans.
XX
PN   WO200037614-A2.
XX
PD   29-JUN-2000.
XX
PF   12-NOV-1999;   99WO-US26704.
XX
PR   18-DEC-1998;   98US-0216295.
XX
PA   (GEMV ) GENENCOR INT INC.

     Mitchinson C,  Wendt DJ;

     WPI; 2000-482483/42.
XX
PT   Novel endoglucanase III or endoglucanase III like cellulase useful for
PT   treating textiles and wood pulp comprises a substitution or deletion at
PT   specified positions in the wild form of endoglucanase III
XX
PS   Example 1; Fig 3; 52pp; English.
XX
CC   The present sequence is a cellulase related to endoglucanase III (EGIII)
CC   from Trichoderma reesei. EGIII like genes were isolated from genomic DNA
CC   libraries constructed from various microorganisms by PCR. The isolated
CC   genes showed significant homology to EGIII from T. reesei. Certain
CC   substitution and deletion mutations have been incorporated into EGIII and
CC   EGIII like cellulases to produce variant enzymes with improved stability,
CC   e.g. increased resistance to temperature stress. The mutants may be used
CC   in textile and wood pulp treatment, as a feed additive, and for reducing
CC   biomass to glucose. They are also useful for stonewashing or indigo dyed
CC   denim and as an agent in laundry and dish detergents.
XX
SQ   Sequence   371 AA;


  Query Match              62.3%;  Score 43;  DB 21;  Length 371;
  Best Local Similarity   66.7%;  Pred. No. 30;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        2 AMGGGRLEGSSC 13
            |   |||| |:|
Db      357 apaggrlnqtsc 368


RESULT   3
AAY84345
ID   AAY84345 standard; Protein; 371 AA.
XX
AC   AAY84345;
XX
DT   12-JUL-2000  (first entry)
XX
DE   Amino acid sequence of an endoglucanase III (EGIII)-like cellulase.
XX
KW   Endoglucanase III; EGIII; EGIII-like cellulase; surfactant stability;
KW   cellulase; textile processing; textile cleaning; stonewashing;
KW   indigo dyed denim; cellulose containing fabric; fabric smoothness;
KW   pill removal; fibril removal; cotton; cellulosic fibre; dying; detergent;
KW   animal feed; wood pulp; paper; grain; biomass reduction; glucose.
XX
OS   Actinomyces sp.
XX
PN   WO200014208-A1.
XX
PD   16-MAR-2000.
XX
PF   24-AUG-1999;   99WO-US19154.
XX
PR   03-SEP-1998;   98US-0146729.
XX
PA   (GEMV ) GENENCOR INT INC.
XX
PI   Fowler T;
XX
DR   WPI; 2000-271052/23.
XX
PT   Novel variant endoglucanase III-like cellulases with improved
PT   surfactant stability and resistance to temperature stress, useful for
PT   textile processing or cleaning, treating wood pulp, food and grain, and
PT   reducing biomass to glucose -
XX
PS   Disclosure; Page 64-65; 73pp; English.
XX
CC   The present sequence represents an endoglucanase III (EGIII)-like
CC   cellulase. The cellulase has homology to the Trichoderma reesei EGIII
CC   protein. The variant cellulases have improved temperature stability,
CC   and improved surfactant stability. The variant cellulases and
CC   compositions containing them are used in textile processing or cleaning,
CC   e.g. stonewashing of indigo dyed denim, and modifying the texture, feel
CC   or appearance of cellulose containing fabrics (e.g. improving fabric
CC   smoothness or removing pills and fibrils). The compositions may also be
CC   used for the removal of immature or dead cotton from cellulosic fibres
CC   or fabric, which can cause uneven dying. The cellulase may also be used
CC   in a detergent composition for washing laundry and dishes and in the
CC   treatment of animal feed, wood pulp, paper, non-animal foods and grains.
CC   The enzymes may also be used in the reduction of biomass to glucose.
XX
SQ   Sequence   371 AA;


  Query Match              62.3%;  Score 43;  DB 21;  Length 371;
  Best Local Similarity   66.7%;  Pred. No. 30;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        2 AMGGGRLEGSSC 13
```

```
          |  |||| |:||
Db    357 apaqgrlnqtsc 368


RESULT   4
AAY67497
ID   AAY67497 standard; Protein; 386 AA.
XX
AC   AAY67497;
XX
DT   19-MAY-2000  (first entry)
XX
DE   Protien sequence of CelA and cellulase 11AG8 fusion.
XX
 W   Cellulase; Actinomycetes; detergent; feed additive; textile treatment;
     pulp; paper; cellulase 11AG8.

     Streptomyces lividans.
OS   Synthetic.
XX
FH   Key                 Location/Qualifiers
FT   Peptide             1..46
FT                       /note= "celA signal sequence"
FT   Protein             47..386
FT                       /note= "cellulase 11AG8 mature peptide"
XX
PN   WO200009707-A1.
XX
PD   24-FEB-2000.
XX
PF   28-MAY-1999;   99WO-US11971.
XX
PR   24-JUN-1998;   98US-0104308.
PR   18-NOV-1998;   98WO-US24649.
PR   28-MAY-1999;   99US-0321981.
XX
PA   (GEMV ) GENENCOR INT INC.
XX
PI   Jones BE,  Van Der Kleij WAH,  Van Solingen P,  Weyler W;
XX
DR   WPI; 2000-224344/19.
DR   N-PSDB; AAZ57031.
XX
PT   A novel Actinomocyes cellulase and related DNA, useful for detergent
PT   compositions, treating textiles and paper or pulp -
XX
     Example 6; Fig 15; 72pp; English.

     The invention provides a cellulase from Actinomycetes. The cellulase can
CC   be used in a detergent composition, as an additive for animal feed and
CC   for the treatment of textiles or pulp and paper. The DNA encoding the
CC   cellulase can be used to identify homologous cellulases and for
CC   recombinant production of cellulases. The present sequence represents
CC   the protein sequence of a celA signal sequence and cellulase 11AG8
CC   fusion sequence contained in the expression cassette consisting of the
CC   GI promoter, celA signal sequence, cellulase 11AG8 and GI terminator
XX
SQ   Sequence   386 AA;


  Query Match             62.3%;  Score 43;  DB 21;  Length 386;
  Best Local Similarity   66.7%;  Pred. No. 31;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AMGGGRLEGSSC 13
                |  |||| |: |
Db        372 apaqgrlnqtsc 383


RESULT   5
AAB39255
ID   AAB39255 standard; Protein; 33 AA.
XX
AC   AAB39255;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Gene 17 human secreted protein homologous amino acid sequence #13[illegible]
XX
KW   Human; secreted protein; immunosuppressive; [illegible]; antirheumatic;
KW   antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW   nootropic; neuroprotective; antibacterial; virucide; fungicide; neoplasm;
KW   opthalmological; autoimmune disease; rheumatoid arthritis; angiogenesis;
KW   hyperproliferative disorder; cardiovascular disorder; infection;
KW   cerebrovascular disorder; nervous system disorder; ocular disorder;
KW   wound healing; chemotaxis.
XX
OS   Mus musculus.
XX
PN   WO200056754-A1.
XX
PD   28-SEP-2000.
XX
PF   16-MAR-2000; 2000WO-US06[illegible].
XX
PR   19-MAR-1999;   99US-0[illegible].
PR   10-DEC-1999;   99US-[illegible]
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen GA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-579433/54.
XX
PT   Isolated nucleic acid molecule encoding a human secreted protein [illegible]
PT   used in preventing, treating or ameliorating a medical condition
XX
PS   Disclosure; Page 34; 434pp; English.
XX
CC   The polynucleotide sequences given in AA[illegible]42[illegible]74[illegible] encode the human
CC   secreted proteins represented in AAB[illegible] sequences
CC   AAB39227-B39308 are alternative proteins encoded by the genes, and its
CC   protein sequences with which they share homology. The proteins have
CC   activities based on the tissues and cells in which they are expressed.
CC   Examples of activities include immunosuppressive; antiarthritic;
CC   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
CC   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
CC   fungicide; and opthalmological. The human secreted proteins,
CC   polynucleotides, antagonists and agonists of the invention may be useful
CC   in the treatment, prevention, and/or diagnosis of various disease,
CC   disorders and conditions such as autoimmune diseases e.g. rheumatoid
CC   arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC   liver, cardiovascular disorders e.g. [illegible] cerebrovascular
CC   disorders e.g. cerebral ischaemia, [illegible] nervous system disorders
CC   e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC   fungi and ocular disorders e.g. corneal infection. The polypeptides can
CC   also be used to aid wound healing and epithelial cell proliferation, to
CC   regenerate tissues, maintain organs before transplantation, in
CC   chemotaxis and as a food additive or preservative e.g. to increase
CC   storage capabilities. Sequences AA[illegible]74[illegible] and AAB[illegible]178 are used
CC   during the isolation and characterisation of the genes of the invention.
XX
SQ   Sequence   33 AA


  Query Match             58.0%;  Score 40;  [illegible]  Length [illegible]
  Best Local Similarity   70.0%;  Pred. No [illegible]
  Matches    7;  Conservative    0;  Mismatches    [illegible]  Indels    0;  Gaps    0;

Qy          4 GGGRLEGSSC 13
              | |: ||:
Db         15 qsqreqsqe 24


RESULT   6
AAG08827
```

```
ID   AAG08827 standard; Protein; 54 AA.
XX
AC   AAG08827;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 6516.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
[illegible]   25-FEB-1999;   99US-0121825.
[illegible]   05-MAR-1999;   99US-0123180.
[illegible]   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
[illegible]   24-MAY-1999;   99US-0135629.
[illegible]   25-MAY-1999;   99US-0136021.
[illegible]   27-MAY-1999;   99US-0136392.
[illegible]   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
```

```
PR    23-AUG-1999;   99US-0149930.
PR    25-AUG-1999;   99US-0150566.
PR    26-AUG-1999;   99US-0150884.
PR    27-AUG-1999;   99US-0151065.
PR    27-AUG-1999;   99US-0151066.
PR    27-AUG-1999;   99US-0151080.
PR    30-AUG-1999;   99US-0151303.
PR    31-AUG-1999;   99US-0151438.
PR    01-SEP-1999;   99US-0151930.
PR    07-SEP-1999;   99US-0152363.
PR    10-SEP-1999;   99US-0153070.
PR    13-SEP-1999;   99US-0153758.
PR    15-SEP-1999;   99US-0154018.
PR    16-SEP-1999;   99US-0154039.
PR    20-SEP-1999;   99US-0154779.
PR    22-SEP-1999;   99US-0155135.
PR    23-SEP-1999;   99US-0155486.
PR    24-SEP-1999;   99US-0155659.
PR    28-SEP-1999;   99US-0156458.
PR    29-SEP-1999;   99US-0156596.
PR    04-OCT-1999;   99US-0157117.
PR    05-OCT-1999;   99US-0157753.
PR    06-OCT-1999;   99US-0157865.
PR    07-OCT-1999;   99US-0158029.
PR    08-OCT-1999;   99US-0158232.
PR    12-OCT-1999;   99US-0158369.
PR    13-OCT-1999;   99US-0159293.
PR    13-OCT-1999;   99US-0159294.
PR    13-OCT-1999;   99US-0159295.
PR    14-OCT-1999;   99US-0159329.
PR    14-OCT-1999;   99US-0159330.
PR    14-OCT-1999;   99US-0159331.
PR    14-OCT-1999;   99US-0159637.
PR    14-OCT-1999;   99US-0159638.
PR    18-OCT-1999;   99US-0159584.
PR    21-OCT-1999;   99US-0160741.
PR    21-OCT-1999;   99US-0160767.
PR    21-OCT-1999;   99US-0160768.
PR    21-OCT-1999;   99US-0160770.
PR    21-OCT-1999;   99US-0160814.
PR    21-OCT-1999;   99US-0160815.
PR    22-OCT-1999;   99US-0160980.
PR    22-OCT-1999;   99US-0160981.
PR    22-OCT-1999;   99US-0160989.
PR    25-OCT-1999;   99US-0161404.
PR    25-OCT-1999;   99US-0161405.
PR    25-OCT-1999;   99US-0161406.
PR    26-OCT-1999;   99US-0161359.
PR    26-OCT-1999;   99US-0161360.
PR    26-OCT-1999;   99US-0161361.
PR    28-OCT-1999;   99US-0161920.
PR    28-OCT-1999;   99US-0161992.
PR    28-OCT-1999;   99US-0161993.
PR    29-OCT-1999;   99US-0162142.

  Query Match             58.0%;  Score 40;  DB 21;  Length 54;
  Best Local Similarity   77.8%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 GGRLEGSSC 13
              ||| ||| |
Db         46 qgrteqsdc 54


RESULT 7
AAW88236
ID   AAW88236 standard; Protein; 432 AA.
XX
AC   AAW88236;
XX
DT   15-MAR-1999  (first entry)
XX
DE   Mouse prothrombinase Fgl2 protein.
XX
KW   Prothrombinase; Fgl2; mouse; immune coagulation; antibody;
KW   inhibitor; infection; graft rejection; glomerulonephritis; cancer;
KW   gastrointestinal disease; foetal loss; therapy; vaccine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   172..174
FT                   /label= Asn is N-glycos. site
FT   Modified-site   228..230
FT                   /note= "Asn is N-glycos. site
FT   Modified-site   257..259
FT                   /note= "Asn is N-glycos. site
FT   Modified-site   323..325
FT                   /note= "Asn is N-glycos. site
FT   Domain          213..419
FT                   /note= "fibrinogen related domain"
XX
PN   WO9851335-A1.
XX
PD   19-NOV-1998.
XX
PF   15-MAY-1998;   98WO-CA00475.
XX
PR   10-OCT-1997;   97US-0061684.
PR   15-MAY-1997;   97US-0046567.
XX
PA   (LEVY/) LEVY G.
XX
PI   Levy G;
XX
DR   WPI; 1999-059687/05.
DR   N-PSDB; AAV84140.
XX
PT   Modulating immune coagulation activity, e.g. using Fgl2 antibodies and
PT   compounds, used to treat conditions including graft rejection and
PT   foetal loss
XX
PS   Claim 8; Page 70-71; 135pp; English.
XX
CC   This is the amino acid sequence of mouse prothrombinase Fgl2, as
CC   predicted from Fgl2 DNA (see AAV84140). Fgl2 is a 70 kDa
CC   transmembrane serine protease that has immune procoagulant activity.
CC   The human Fgl2 amino acid sequence is given in AAW88237. The
CC   invention provides a method for inhibiting immune coagulation by
CC   inhibiting the activity or expression of Fgl2. The method can be
CC   used in vivo to treat a condition which requires a reduction in
CC   immune coagulation such as bacterial and viral infections, cancer
CC   glomerulonephritis, a number of gastrointestinal diseases
CC   allograft and xenograft rejection and foetal loss. An Fgl2-specific
CC   antibody, an Fgl2 antisense oligonucleotide or a substance that
CC   affects prothrombinase activity of a Fgl2 protein may be used to
CC   treat a condition requiring a reduction in coagulant activity
CC   A vaccine containing an Fgl2 protein or peptide is used for
CC   prevention of graft rejection or foetal loss (claimed)
XX
SQ   Sequence   432 AA;


  Query Match             58.0%;  Score 40;  DB [illegible];  Length 432;
  Best Local Similarity   70.0%;  Pred. No. [illegible];
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GGGRLEGSSC 13
              | || |||  |
Db         44 qsqrceqsqc 53


RESULT 8
AAY36511
ID   AAY36511 standard; Protein; 25 AA.
XX
```

```
AC   AAY36513;
XX
DT   17-SEP-1999  (first entry)
XX
DE   Fragment of human secreted protein encoded by gene 32.
XX
KW   Human; secreted protein; cancer; tumour; developmental abnormality;
KW   foetal deficiency; blood disorder; immune system disorder; inflammation;
KW   autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW   schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW   atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW   digestive disorder; endocrine disorder; infection; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO9931117-A1.
XX
PD   24-JUN-1999.
XX
     17-DEC-1998;   98WO-US27059.

     19-DEC-1997;   97US-0068369.
     18-DEC-1997;   97US-0068006.
PR   18-DEC-1997;   97US-0068007.
PR   18-DEC-1997;   97US-0068008.
PR   18-DEC-1997;   97US-0068053.
PR   18-DEC-1997;   97US-0068054.
PR   18-DEC-1997;   97US-0068057.
PR   18-DEC-1997;   97US-0068064.
PR   18-DEC-1997;   97US-0070923.
PR   19-DEC-1997;   97US-0068169.
PR   19-DEC-1997;   97US-0068365.
PR   19-DEC-1997;   97US-0068367.
PR   19-DEC-1997;   97US-0068368.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Carter KC,  Duan RD,  Feng P,  Ferrie AM,  Florence C;
PI   Florence K,  Greene JM,  Janat F,  Kyaw H,  Moore PA;
PI   Ni J,  Rosen CA,  Ruben SM,  Shi Y,  Soppet DR,  Wei Y;
PI   Yu G;
XX
DR   WPI; 1999-418749/35.
XX
PT   New isolated human genes encoding secreted polypeptides
XX
PS   Disclosure; Page 469; 537pp; English.
XX
CC   AAX97907 to AAX98029 represent 110 isolated human secreted protein
CC   genes. AAY36224 to AAY36727 represent the secreted proteins encoded by
CC   the 110 human genes. The genes and their corresponding secreted
     polypeptides are useful for preventing, treating or ameliorating medical
     conditions, e.g. by protein or gene therapy. Also pathological conditions
     can be diagnosed by determining the amount of the new polypeptides in a
     sample or by determining the presence of mutations in the new genes.
CC   Specific uses are described for each of the 110 genes, based on which
CC   tissues they are most highly expressed in, and include developing
CC   products for the diagnosis or treatment of cancer, tumours, developmental
CC   abnormalities and foetal deficiencies, blood disorders, diseases of the
CC   immune system, autoimmune diseases, inflammation, allergies, Alzheimer's
CC   and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,
CC   sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular
CC   disorders, kidney disorders, digestive/endocrine disorders, infections
CC   and AIDS. The polypeptides are also useful for identifying their binding
CC   partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are
CC   used in the exemplification of the present invention.
XX
SQ   Sequence   25 AA;

  Query Match             56.5%;  Score 39;  DB 20;  Length 25;
  Best Local Similarity   70.0%;  Pred. No. 10;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          4 GGGRLEGSSC 13
                |||: || ||
Db          2 qqqkeeqvsc 11


RESULT   9
AAB95038
ID   AAB95038 standard; Protein; 106 AA.
XX
AC   AAB95038;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:16795.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116125.
XX
PR   29-JUL-1999;    99JP-0248036.
PR   27-AUG-1999;    99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs
XX
PS   Claim 8; SEQ ID 16795; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5' end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5' end sequence/3' end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   106 AA;


  Query Match             56.5%;  Score 39;  DB 22;  Length 106;
```

```
  Best Local Similarity   70.0%;  Pred. No. 39;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 GGGRLEGSSC 13
           |||| ||  |
Db      11 gqgrwegvgc 20


RESULT  10
AAY97667
ID   AAY97667 standard; Protein; 647 AA.
XX
AC   AAY97667;
XX
 T   08-MAY-2001  (first entry)

     Zea mays ZmEIN3 2 protein sequence.

KW   Zm EIN3-1; Zm EIN3-2; ein; ethylene signal transduction pathway; maize;
KW   ethylene-mediated response; crowding tolerance; seed set; maturation;
KW   seed development; growth in impacted soil; flooding tolerance;
KW   senescence; disease resistance.
XX
OS   Zea mays.
XX
PN   WO200105953-A2.
XX
PD   25-JAN-2001.
XX
PF   07-JUL-2000; 2000WO-US18746.
XX
PR   15-JUL-1999;   99US-0143965.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Simmons CR;
XX
DR   WPI; 2001-147338/15.
DR   N-PSDB; AAA91185.
XX
PT   Novel maize ethylene signalling pathway EIN3 gene useful for modulating
PT   the level of EIN3 in maize plants, including crowding tolerance, growth
PT   in impacted soils, flooding tolerance and disease resistance  -
XX
PS   Claim 10; Page 82-85; 86pp; English.
XX
 C   This sequence is the corn Zm EIN3 2 protein. The protein is involved
     in the ethylene signal transduction pathway, and is an EIN3 homologue.
 C   An expression cassette comprising the DNA sequence is useful for
CC   modulating the level of EIN3 in a plant, in particular a maize plant. The
CC   maize genes are nuclear transcription factors that promote the
CC   ethylene-mediated responses, including crowding tolerance, seed set and
CC   development, growth in impacted soils, flooding tolerance, maturation and
CC   senescence and disease resistance. Diminishment of ethylene action in
CC   plant, in particular cereals such as maize, by reducing the expression or
CC   activity of the DNA promotes tolerance to close spacing with reduced
CC   stress and yield loss. The DNA is useful as a probe or amplification
CC   primer in the detection, quantitation or isolation of gene transcripts,
CC   in detecting deficiencies in the level of mRNA in screening for desired
CC   transgenic plants, for detecting mutations in gene, for monitoring
CC   upregulation of expression or changes in enzyme activity in screening
CC   assays, orthologs, or paralogs of the gene, or for site directed
CC   mutagenesis in eukaryotic cells. The nucleic acid can also be used for
CC   recombinant expression of polypeptides or as immunogens in the
CC   preparation and/or screening of antibodies. The proteins can be employed
CC   in assays for enzyme agonists or antagonists of enzyme function, or as
CC   immunogens or antigens to obtain antibodies specifically immunoreactive
CC   with the protein. Plants expressing the DNA germinate better in compacted
CC   soils and in flooded conditions or water-logged soils, resulting in
CC   higher stand counts.
XX
SQ   Sequence   647 AA;

  Query Match             56.5%;  Score 39;  DB 2;  Length 647;
  Best Local Similarity   66.7%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MAMCGGRLEGSS 12
           || |||  :|||
Db     106 masqqqladgss 117


Search completed: March 14, 2002 17:2[illegible]
Job time: 7904 sec
```